

Non-Parametric 3D Shape Warping

Ulrich Hillenbrand

*Institute of Robotics and Mechatronics, German Aerospace Center (DLR), D-82234 Wessling
Ulrich.Hillenbrand@dlr.de*

Abstract—A method is presented for non-rigid alignment of a source shape to a target shape through estimating and interpolating point-wise correspondences between their surfaces given as point clouds. The resulting mapping can be non-smooth and non-isometric, relate shapes across large variations, and find partial matches. It does not require a parametric model or a prior of deformations. Results are shown for some objects from the Princeton Shape Benchmark and a range scan.

I. INTRODUCTION

Shape warping is the process of estimating a correspondence or mapping between the points defining two shapes of the same or a similar category. Applications arise, e.g., when tracking shape deformations, comparing shapes for locating and quantifying similarity, and building shape category models from examples. Especially for the latter two cases, warping techniques are desirable that cover large and diverse shape variations without prior knowledge.

Most previous work on warping of 3D surfaces has been in the context of registering range data sets while compensating for low-frequency errors or modeling of a deformable object. Some of those methods have been based on non-rigid extensions to the iterative closest point (ICP) algorithm [1], [2], [3]. While in ICP-type algorithms correspondences and alignment are estimated in an alternating fashion with separate cost functions, an alternative technique optimizes both correspondences and alignment simultaneously in a joint cost function [4]. Approximately isometric (in the geodesic sense) deformations can be estimated by a variant of multidimensional scaling [5].

One problem with all these approaches to shape warping is that they need regularization by constraints or penalty terms, enforcing some degree of smoothness, stiffness, rigidity, or isometry, which effectively act as a deformation prior. While this makes sense in the application contexts of the above methods, for modeling shape variability within an object category, such constraints are generally not valid. Another drawback is their formulation as a non-convex optimization problem, such that convergence to the global optimum is hard to ensure.

The algorithm described in this paper differs from previous ones for 3D shape warping in two important respects.

- It is not formulated as a (non-convex) optimization problem, but rather as one of Monte-Carlo-style density estimation, avoiding most problems with global convergence, and thus handling large misalignment and deformations between shapes.
- It does not use a probabilistic model or prior of deformations. In this sense, the estimator is non-parametric and

purely data-driven, and thus handles even non-smooth and non-isometric deformations.¹

Like some of the other techniques, it finds a partial match between shapes that are not fully consistent.

II. THE ALGORITHM

The three main steps of the shape warping algorithm are

- A. deformation-tolerant pose estimation,
- B. correspondence estimation,
- C. surface-point mapping.

It is worth noting that in this procedure, there is no joint estimation of alignment and correspondences, in contrast to most other schemes. Rather, the initial rigid alignment is estimated without determining correspondences. Once a reasonable rigid alignment is achieved, estimation of correspondences is a comparably simple step. In turn, once correspondences are determined, mapping of surface points is trivial. The most critical step in the proposed procedure is hence the initial rigid alignment.

Each of these steps will now be explained. The warping is directed from a *source shape* to a *target shape*.

A. Deformation-tolerant pose estimation

The algorithm starts by globally estimating a rigid alignment, or pose, of the source shape to the target shape. The rigid alignment has to tolerate significant deviations of the target shape from the source shape. A reasonable alignment surely is one where corresponding parts of the two objects come as close as possible to each other. This not only reflects our intuition about the ‘right’ alignment of two different shapes, it is also crucial for the following step of correspondence estimation. However, we seek such an alignment without a hint as to the correct correspondences.

A correspondence-free alignment that is also robust to large geometric deviations is provided within the framework of parameter-density estimation and maximization, or parameter clustering. This is a robust estimation technique based on location statistics in a parameter space where parameter samples are computed from data samples [6], [7]. The estimator may be viewed as a continuous version of a generalized, randomized Hough transform.² In the present variant, the sampling is from a surface description based on points with their local surface

¹Note, however, that even non-parametric techniques necessarily depend upon the choice of algorithmic parameters.

²Parameter clustering is fundamentally different from RANSAC-style techniques, in that no quality measure of hypotheses, or any kind of statistics, is evaluated in data space.

normal vector, which we shall refer to as *surflets*. In particular, no special geometric features or high-level primitives need to be extracted from the data, making the procedure applicable to dense range data of all shapes.

The surface normals can be estimated from range data by local covariance statistics: the eigenvector for the smallest eigenvalue of the covariance matrix, computed from points within a local surface region, is a useful estimate of surface normal orientation. The outward direction of the normal is known from the gaze direction of the sensor.

Let $\mathcal{S} \subset \mathbb{R}^3$ be the given set of points on the source shape and $\mathcal{T} \subset \mathbb{R}^3$ the given set of points on the target shape. Let further be $u(s)$ and $v(t)$ the (normalized) surface normal vectors for points $s \in \mathcal{S}$ and $t \in \mathcal{T}$, respectively.

A pose hypothesis can now be computed from a minimum subset of two source surflets matched against a minimum subset of two target surflets. The sampling proceeds thus as follows.

- 1) Randomly draw a point pair $s_1, s_2 \in \mathcal{S}$.
- 2) Randomly draw a point pair $t_1, t_2 \in \mathcal{T}$, such that the surflet pair $\{(t_1, v(t_1)), (t_2, v(t_2))\}$ is geometrically similar to $\{(s_1, u(s_1)), (s_2, u(s_2))\}$.
- 3) Compute a rigid motion that aligns $\{(s_1, u(s_1)), (s_2, u(s_2))\}$ to $\{(t_1, v(t_1)), (t_2, v(t_2))\}$, up to distortions.
- 4) Compute and store the six parameters of the hypothetical motion.

In step 2) of the sampling procedure, a metric for surflet pairs and an efficient procedure for enforcing similarity when drawing from \mathcal{T} are needed. The intrinsic geometry of a surflet pair can be smoothly described by four parameters, in various ways. The Euclidean metric in such a parameter space will provide a similarity measure for surflet pairs. Similarity of the surflet pair sampled from \mathcal{T} in step 2) to the one sampled from \mathcal{S} in step 1) is efficiently enforced by indexing into a hash table of surflet pairs previously sampled from \mathcal{T} . The table is accessed through the four parameters of the drawn \mathcal{S} -pair as the key.

Rigidly aligning a surflet pair with another in step 3) requires trading off between positional and directional information. Unlike for pure point sets, there is no unique principled formulation of a cost function. Here we estimated the rotation from the surface normals alone, while the translation has to be estimated from the surface points. More precisely, the rotation between the two surflet pairs $\{(s_1, u(s_1)), (s_2, u(s_2))\}$ and $\{(t_1, v(t_1)), (t_2, v(t_2))\}$ was computed to minimize the squared angles between the normals, i.e.,

$$R = \arg \min_{R' \in SO(3)} \sum_{i=1}^2 \arccos^2(v(t_i) \cdot R' u(s_i)) , \quad (1)$$

and the translation is then the least-squares solution on the points, i.e.,

$$T = \arg \min_{T' \in \mathbb{R}^3} \sum_{i=1}^2 \|R s_i + T' - t_i\|^2 . \quad (2)$$

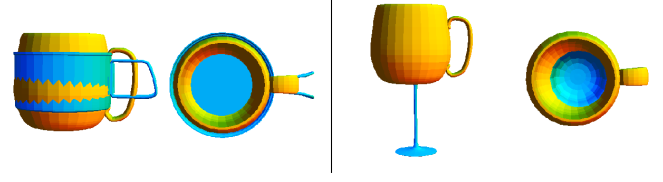


Fig. 1. Alignment found for two different mugs, and for one of the mugs and a liqueur glass.

The parameterization of rigid motions chosen for sampling step 4) may have an influence on the result. In fact, the parameter density from which we sample depends upon this choice. A parameterization that is consistent for clustering is generally recommended, in the sense of [6].

By repeatedly executing the sampling procedure 1)–4) above, we obtain samples from the parameter density for the rigid alignment problem. This parameter density is similar in spirit to a posterior density, but without assuming a probabilistic observation model [6].

The parameter samples can be stored in an array or a tree of bins. The sampling stops, when a significant cluster of samples has formed, as judged from the bin counts. Then the location of maximum parameter density is searched by starting a mean-shift procedure [8], [9] from the center of the bin with the highest parameter count. The mean-shift procedure is repeatedly started from the bin with the next highest parameter count, until the density found in the converged mean-shift window is significantly lower than found in the first run of mean shift. From all the local density maxima found through mean shift, the location in the 6D parameter space of the largest maximum is returned as the pose estimate $(\hat{R}, \hat{T}) \in SO(3) \times \mathbb{R}^3$ of the target shape relative to the source shape. By re-starting the mean shift from several points in high-density regions of parameter space, it is likely that the strongest mode of the density is found. Details of the implementation will be presented elsewhere [10].

Figure 1 shows two examples of alignment of different shapes found through the described procedure.

B. Correspondence estimation

Correspondence is here defined as a directed relation between points of the source shape and points of the target shape, and symmetry shall be enforced only at the mapping stage; see next section. Hence, not all directed correspondences established will affect the final mapping of surface points. In a sense, the true correspondences will only be those realized by the final mapping.

Since rigid alignment of source and target shapes has brought corresponding parts already close to each other, we can again rely on the simple local surface description by surflets to find correspondences, based on proximity of points and alignment of normal vectors.

The surface normal vector $u(s)$ at a source point $s \in \mathcal{S}$ is registered with the target surface through the rotation \hat{R} , the estimate from initial alignment of the shapes; cf. previous

section. We define the set of target points with a normal vector $v(t)$ oriented at most an angle δ away from $\hat{R}u(s)$,

$$\mathcal{T}_\delta^{\hat{R}}(s) = \left\{ t \in \mathcal{T} \mid v(t) \cdot \hat{R}u(s) > \cos \delta \right\}. \quad (3)$$

The points belonging to $\mathcal{T}_\delta^{\hat{R}}(s)$ can be efficiently retrieved from a hash table indexed by normal orientation. Then the *forward correspondence* of point $s \in \mathcal{S}$ is

$$C_f(s) = \arg \min_{t \in \mathcal{T}_\delta^{\hat{R}}(s)} \|\hat{R}s + \hat{T} - t\|. \quad (4)$$

Likewise, we define the set of source points

$$\mathcal{S}_\delta^{\hat{R}}(t) = \left\{ s \in \mathcal{S} \mid v(t) \cdot \hat{R}u(s) > \cos \delta \right\}, \quad (5)$$

and the *backward correspondence* of point $t \in \mathcal{T}$ is

$$C_b(t) = \arg \min_{s \in \mathcal{S}_\delta^{\hat{R}}(t)} \|\hat{R}s + \hat{T} - t\|. \quad (6)$$

The shape warps shown in this paper have used a tolerance parameter $\delta = 35$ degrees for surface normal misalignment.

The constraints on correspondence derived from point distance and inter-normal angle are here enforced in a sequential manner. This avoids trading off Euclidean point differences against angular normal differences between surflets. Alternatively, one could employ a surflet metric that has an Euclidean contribution for the points and an angular contribution for the normals.

C. Surface-point mapping

The mapping process interpolates the previously found directed correspondences and enforces their symmetry. Interpolation was done here in the most primitive way. For every source point queried, the n closest points from \mathcal{S} are found, in the Euclidean metric. The forward-mapped point on the target shape then is the average of the n forward-corresponding target points. Formally, a point x on the source shape (but not necessarily in \mathcal{S}) gets *mapped forward* according to

$$M_f(x) = \frac{1}{n} \sum_{s \in \mathcal{S}_n(x)} C_f(s), \quad (7)$$

where $\mathcal{S}_n(x) \subseteq \mathcal{S}$ is the set of n nearest neighbors of x within \mathcal{S} . Likewise, a point y on the target shape (but not necessarily in \mathcal{T}) gets *mapped backward* according to

$$M_b(y) = \frac{1}{n} \sum_{t \in \mathcal{T}_n(y)} C_b(t), \quad (8)$$

where $\mathcal{T}_n(y) \subseteq \mathcal{T}$ is the set of n nearest neighbors of y within \mathcal{T} . The shape warps shown in this paper have used a neighborhood of $n = 3$ for interpolation. Equations (7) and (8) represent something like a zeroth-order interpolation of a discrete map. Alternatively, more refined methods of interpolation between correspondences can be tried, such as distance-weighted averaging, higher-order splines, or thin-plate splines.

Symmetry of interpolated correspondences is enforced through a forward/backward consistency check of the mappings. A point x on the source shape is mapped to a point

$y = M(x) := M_f(x)$ on the target shape, if and only if its forward and backward mappings agree within a radius of ϵ , that is,

$$\|x - M_b(M_f(x))\| < \epsilon. \quad (9)$$

Note that the forward/backward consistency check is similar in spirit to the left/right consistency check in stereo image processing: it eliminates false matches from the map, allowing for partial matches between very different shapes. The shape warps shown in this paper have used a tolerance parameter $\epsilon = 0.04$ object-bounding-box diagonals for forward/backward mapping discrepancy. The result of mapping is the final result of warping.

III. RESULTS

In this section, the behavior of the shape warping algorithm is illustrated by some representative processing examples. The test objects, some mugs and a liqueur glass, have been taken from the Princeton Shape Benchmark [11]. Moreover, a scan of a mug with our laser stripe profiler [12] is included as a test case to also demonstrate the applicability to noisy and partial representations.

For visualization of the shape warps, a regular pattern of surface points from the source mug, colored to code the three Cartesian coordinates, is mapped through (7), conditioned on (9), to the target objects; see fig. 2.

Evidently, the warps are fairly smooth and generally connect corresponding points on the object surfaces. In particular, functionally critical parts like handles and mouths are correctly mapped.

In the example shown in fig. 2a), the source and target shapes are fairly similar, resulting in a rather complete mapping of the source surface to the target surface.

In the example shown in fig. 2b), when warping from the source mug to the target mug, a change in the mugs' surface topology occurs: the handle of the target mug consists of two separate handle parts. The shape warp respects this topological change by mapping the two sides of the source handle to the closer of the two parts of the target handle. Only very few points get interpolated in between the two parts.

In the example shown in fig. 2c), the handle of the target mug is degenerated to a tiny extension near the rim of the mouth. As a result, only the upper part of the source handle is mapped to the degenerate target handle.

When warping from the source mug to the liqueur glass as shown in fig. 2d), the corresponding parts of the objects are correctly identified. In particular, it can be inferred from the warp that the liqueur glass has no handle, while the mug has no stem.

When warping from the source mug to the scanned mug as shown in fig. 2e), the fraction of the source mug corresponding to the partial scan is correctly identified; cf. fig. 3 for the scanned mug data.

IV. DISCUSSION

A procedure for establishing meaningful correspondences between points on different shapes from the same or similar

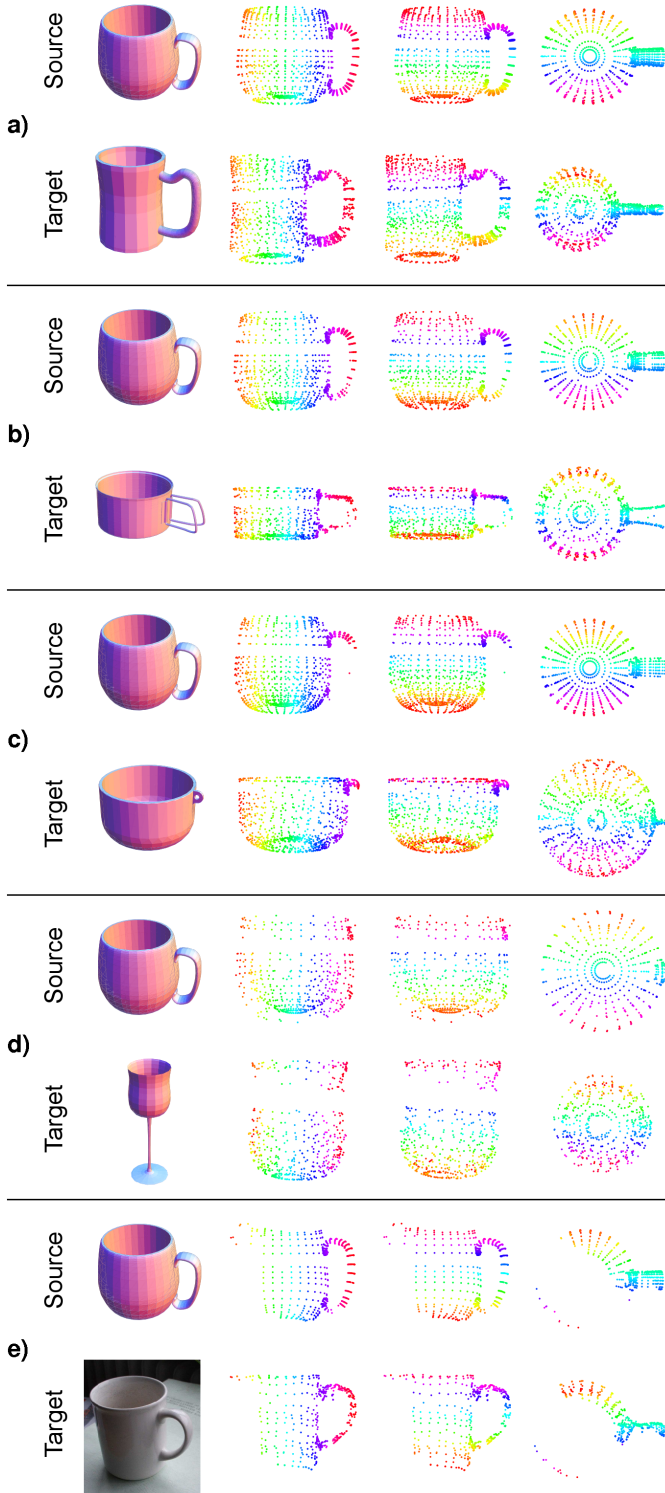


Fig. 2. Shape warping examples. A regular pattern of surface points from the source mug and their mappings to the target shape are colored to code their three Cartesian source coordinates. In each example, only points satisfying condition (9) are shown.

categories has been proposed. The described technique makes no strong assumptions as to the nature of the shape variation.



Fig. 3. Partial scan of a mug used as target shape in fig. 2e).

It should hence be useful in situations where such knowledge is weak, e.g., when comparing unknown shapes for locating and quantifying similarity, or when building shape category models from examples.

Its generality may, on the other hand, also give rise to a potential weakness of the presented method. Thus, there may well be situations where satisfactory point correspondences across shapes need guidance by additional constraints or more specific geometric features than just aligned surface points and normals.

In general, there is no uniquely correct way of making correspondences between points across different shapes. This fact makes it hardly possible to quantify performance of shape warping without a specific application context or other constraints derived from known inter-shape relations.

REFERENCES

- [1] D. Hähnel, S. Thrun, and W. Burgard, "An extension of the ICP algorithm for modeling nonrigid objects with mobile robots," in *Proc. Int. Jt. Conf. Artificial Intelligence*, 2003.
- [2] B. J. Brown and S. Rusinkiewicz, "Non-rigid range-scan alignment using thin-plate splines," in *Proc. Int. Symp. 3D Data Processing, Visualization and Transmission*, 2004.
- [3] B. Amberg, S. Romdhani, and T. Vetter, "Optimal step nonrigid ICP algorithms for surface registration," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2007.
- [4] H. Li, R. W. Sumner, and M. Pauly, "Global correspondence optimization for non-rigid registration of depth scans," in *Proc. Eurograph. Symp. Geometry Processing*, 2008.
- [5] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching," *Proc. Natl. Acad. Sci.*, vol. 103, pp. 1168–1172, 2006.
- [6] U. Hillenbrand, "Consistent parameter clustering: definition and analysis," *Patt. Recogn. Let.*, vol. 28, pp. 1112–1122, 2007.
- [7] —, "Pose clustering from stereo data," in *Proc. VISAPP Int. Workshop on Robotic Perception*, 2008, pp. 23–32.
- [8] K. Fukunaga and L. D. Hostetler, "The estimation of a gradient of a density function, with applications in pattern recognition," *IEEE Trans. Info. Theory*, vol. 21, pp. 32–40, 1975.
- [9] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 24, pp. 603–619, 2002.
- [10] U. Hillenbrand and A. Fuchs, "Four variants of pose clustering from dense range data," 2009, submitted manuscript.
- [11] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton Shape Benchmark," in *Proc. Shape Modeling International*, 2004, <http://shape.cs.princeton.edu/benchmark/>.
- [12] K. H. Strobl, W. Sepp, E. Wahl, T. Bodenmüller, M. Suppa, J. F. Seara, and G. Hirzinger, "The DLR multisensory hand-guided device: the laser stripe profiler," in *Proc. IEEE Int. Conf. on Robotics & Automation*, 2004, pp. 1927–1932.